DOCUMENT RESUME

ED 470 135 IR 021 575

AUTHOR	Ku, Heng-Yu; Sullivan, Howard J.
TITLE	Effects of Personalized Instruction on Mathematics Word Problems in Taiwan.
PUB DATE	2001-11-00
NOTE	12p.; In: Annual Proceedings of Selected Research and Development [and] Practice Papers Presented at the National Convention of the Association for Educational Communications and Technology (24th, Atlanta, GA, November 8-12, 2001). Volumes 1-2; see IR 021 504.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; Foreign Countries; Grade 4; *Individualized Instruction; *Instructional Effectiveness; Intermediate Grades; *Mathematics Achievement; Student Attitudes; Teaching Methods; *Word Problems (Mathematics)
IDENTIFIERS	Taiwan

ABSTRACT

This study investigated the effects of personalized instruction on the achievement and attitudes of Taiwanese students on two-step mathematics word problems. A total of 136 fourth graders in a Taiwanese public school participated in the study. Subjects initially completed a Student Survey on which they chose their favorite foods, sports, stores, classmates, and other selections. The most popular items were then used to create personalized math word problems for the pretest, personalized instructional program, and posttest. Subjects were blocked by ability based on their pretest scores and were randomly assigned within ability levels to either a personalized or non-personalized version of the print-based instructional program. After finishing the program, subjects completed a student attitude survey and the posttest. A repeated-measures ANOVA revealed that subjects in the personalized treatment made significantly greater pretest-to-posttest gains than those in the non-personalized treatment. Subjects also performed significantly better on the personalized pretest and posttest problems than on the non-personalized problems. Personalized subjects and higher-ability students both had significantly more positive attitudes toward the instructional program than their non-personalized and lower-ability counterparts. (Contains 21 references.) (Author/AEF)

Reproductions supplied by EDRS are the best that can be made

ED 470 135

Effects of Personalized Instruction on Mathematics Word Problems in Taiwan

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

P Harris

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

By: Heng-Yu Ku & Howard J. Sullivan

IR021575

BEST COPY AVAILABLE

Effects of Personalized Instruction on Mathematics Word Problems in Taiwan

Heng-Yu Ku
University of Northern Colorado
Howard J. Sullivan
Arizona State University

Abstract

This study investigated the effects of personalized instruction on the achievement and attitudes of Taiwanese students on two-step mathematics word problems. A total of 136 fourth-graders in a Taiwanese public school participated in the study. Subjects initially completed a Student Survey on which they chose their favorite foods, sports, stores, classmates, and other selections. The most popular items were then used to create personalized math word problems for the pretest, personalized instructional program, and posttest. Subjects were blocked by ability based on their pretest scores and were randomly assigned within ability levels to either a personalized or non-personalized version of the print-based instructional program. After finishing the program, subjects completed a student attitude survey and the posttest. A repeated-measures ANOVA revealed that subjects in the personalized treatment made significantly greater pretest-to-posttest gains than those in the non-personalized treatment. Subjects also performed significantly better on the personalized pretest and posttest problems than on the non-personalized problems. Personalized subjects and higher-ability students both had significantly more positive attitudes toward the instructional program than their non-personalized and lower-ability counterparts.

Introduction

Data from the National Assessment of Educational Progress (1992a, 1992b) indicate that mathematics word problems are difficult for students at all age levels in elementary and secondary schools. A major cause of the difficulty appears to be the students' inability to convert the problems into the math operations that must be performed to solve them (Hart, 1996). Some researchers have also noted that lack of familiarity with word problem structures may also contribute to poor student performance (Mayer, 1982; Rosen, 1984).

Personalizing mathematics word problems, such as incorporating personal background information into the problem content, can lead to improvements in performance (Anand & Ross, 1987; Davis-Dorsey, Ross, & Morrison, 1991; Lopez & Sullivan, 1991, 1992). Anand and Ross (1987) tested the effect of using computer-assisted instruction to personalize mathematics instruction for elementary school children. Students who received personalized instruction scored significantly higher on math word problems involving rule recognition and transfer than those whose instruction was not personalized. The authors claimed that personalized contexts increased students' comprehension and motivation by helping them interpret important information in the problem statement. Davis-Dorsey et al. (1991) found that both second-grade and fifth-grade students made significant achievement gains from rewording and personalization of the context in math word problems. In two separate studies with rural eighth-grade Hispanic American students, Lopez and Sullivan (1991, 1992) found significant overall achievement effects for personalization on one-step and two-step mathematics word problems.

Several studies have found that student attitudes are more positive when student interests and preferences are incorporated into instruction in order to personalize it. Ross and his colleagues (Ross, 1983; Ross, McCormick, & Krisak, 1986; Ross, McCormick, Krisak, & Anand, 1985) employed personalization in a series of adaptive instruction studies. Favorable attitude results were obtained when preservice teachers received education-related materials and nursing students received medical-related materials, and poorer results were attained when each group received the other's materials. Herndon (1987) found that high school students who received instruction based on common group interests had significantly more favorable attitudes and higher return-to-task motivation than students whose instruction was not interest-based. Cordova and Lepper (1996) and Hart (1996) also found more favorable attitudes or motivation toward personalized than toward non-personalized instruction.

Researchers have offered two theory-based explanations for the effectiveness of personalized instruction in studies where it has yielded better results than non-personalization. One is that students' greater familiarity with personalized problem situations and content may enable them to solve problems more easily by reducing their cognitive load (Lopez & Sullivan, 1991, 1992; Miller & Kulhavy, 1991). This position is supported by D'Ailly, Simpson, and MacKinnon's (1997) statement that "self-referencing facilitates general encoding processes and decreases the load on working memory during problem solving."

The second explanation is based on interest theory (Mayer, 1998). Mayer notes that students exert more effort and are more successful solving problems that interest them than problems that do not. Several researchers have cited greater student interest and motivation as reasons for better performance under personalized instruction (Cordova & Lepper, 1996; Lopez & Sullivan, 1992; Ross & Anand, 1987). The "reduced cognitive load" and "increased interest" explanations appear to be compatible with one another rather than being alternative or competing explanations.

The present study was designed to address the unanswered questions from the Ku and Sullivan (2000) research and to investigate the stability of the other findings from that study and other research. The likelihood of a ceiling effect, which confounded interpretation of the personalization versus non-personalization issue in the earlier research, was reduced by conducting this study with fourth-grade students instead of fifth graders as in the earlier study. This study was also conducted with 136 students, or nearly twice as many as the earlier one, to increase the power of the statistical analyses. These changes were designed to provide a clearer answer to the question of whether instruction incorporating personalized math word problems would yield significantly better student performance than instruction using only non-personalized items.

The question of whether merely stating test items in personalized form would yield better performance than stating them in non-personalized form, irrespective of the type of instruction students receive, was addressed by including an equal number of personalized and non-personalized items on the pretest, as well as on the posttest. Including personalized and non-personalized items on the pretest, as well as on the posttest, in the present study was designed to permit more direct and stable comparisons of the effects of the two types of test items alone and in combination with the instructional treatments.

The present study investigated the effects of two levels of group personalization (personalized, non-personalized) on the achievement of fourth-grade Taiwanese students on two-step mathematics word problems. Mathematics ability level, as determined by scores on the pretest administered prior to the instructional phase of the study, was also included as a variable because it is an important factor in mathematics achievement and because of differential findings by ability level in earlier research on personalization (Dwyer, 1996; Ku & Sullivan, 2000; Lopez & Sullivan, 1991).

The primary research questions for the study were as follows:

1. Does personalization of instruction increase the achievement of Taiwanese students on mathematics word problems?
2. Does personalization of instruction have a differential effect on the performance of higher-ability and lower-ability students on mathematics word problems?
3. Do Taiwanese students perform better on personalized word problems than on non-personalized problems irrespective of the type of instruction (personalized or non-personalized) they receive?
4. Does personalization of mathematics word problems influence student attitudes toward the instruction on these problems?

Method

Subjects

A total of 136 fourth-grade Taiwanese students from four classes taught by different teachers at a public elementary school in Taiwan participated in this study. The school is located in a mid-level income and socioeconomic area in Fengyuan, a city with a population of approximately 160,000 people.

Materials

Student Survey. A 20-item student survey was used to determine the personal backgrounds and interests of the participants. Topics included the names of the students' favorite places, activities, sports, friends, convenience stores, foods, and so forth. Students wrote in two favorite responses for each survey item.

The survey was administered one week prior to the pretest. Responses to each survey item were tabulated by the experimenter and then used to design the personalized version of the instructional program and the tests.

Instructional Program. Two parallel versions of an instructional program on two-step math word problems were developed in print form in Chinese. Taiwanese students learn addition and subtraction in the third grade and multiplication and division in the fourth grade. The word problems in the instructional program and the test items were taken directly from the fourth-grade and fifth-grade mathematics

textbooks used by the participants. The program was administered after the students had studied one-step multiplication and division, and it covered two-step processes that they had not yet formally studied.

The non-personalized version of the instructional program was written first and included standard problem types from the students' math textbook. The personalized version was then written by incorporating the most popular referents (places, foods, sports, names, etc.) from the Student Survey into the previously non-personalized version.

An example of a practice item in personalized and non-personalized form is provided below. The example item requires using multiplication followed by division.

Non-personalized:	The teacher has 2 dozen cans of soft drink to be shared equally by 8 students. How many cans of soft drink does each student get?
Personalized:	The teacher, Ms. Sue, has 2 dozen cans of milk tea to be shared equally by 8 students. How many cans of milk tea does each student get?

The instructional program covered procedures for solving two-step word problems involving four different combinations of multiplication and division operations (multiply-multiply, multiply-divide, divide-multiply, divide-divide) with whole numbers. A four-step strategy based on the work of Enright and Choate (1993) was incorporated into the instructional program for both treatments. The four steps, which had not been taught to the students previously, were: 1. Read the question. 2. Think through the problem. What must be found out? What steps are involved? 3. Choose the steps and do the math. 4. Check your answer.

Instruction for solving each of the four types of problems contained the four-part Enright and Choate strategy with two worked examples for each problem type. After the examples for each problem type, the instructional program contained two practice problems for the students to work. Each pair of practice problems was placed on a page with enough open space to allow students to work out the problems. When students completed the instructional program, the experimenter collected the programs and scored the practice problems. On the following day, the experimenter conducted a review in which he provided the answers to the practice problems and the explanations for them. All materials used in the study were in Chinese.

Procedures

The Student Survey was administered to participants two weeks prior to the treatment. The most popular choices (places, foods, sports, and so forth) from the survey were subsequently used to convert the non-personalized problems into the personalized content for the pretest, the instructional program, and the posttest. The pretest was administered one week after students filled out the Student Survey.

After the pretests had been scored, the subjects were blocked within each class by their pretest scores into higher-ability and lower-ability groups, and were assigned within blocks to either the personalized or the non-personalized version of the instructional program. Sixty-eight subjects each, 34 higher-ability and 34 lower-ability, were in the personalized treatment and the non-personalized treatment. The mean pretest score was 9.35 (SD = 1.65) for the 68 higher-ability subjects and 3.74 (SD = 2.11) for the 68 lower-ability subjects.

The experimental part of the study took place over three 40-minute class periods on three different days one week after the pretest. The experimenter served as the instructor for all treatments in regularly scheduled math classes in four different classrooms. The experimenter read instructions to all students and told them that they would be helping with the development of a new math program in Taiwan and that they should try their best to learn and to solve the problems. On the first day, subjects completed the instructional part of the program. On day two, the experimenter went over the eight practice problems with subjects and wrote the answers on the blackboard. After this review, students filled out the student attitude survey and teachers completed the teacher attitude survey. On the final day, subjects took the posttest.

Criterion Measures

Pretest. A total of 24 problems were developed, in both personalized and non-personalized forms, for the pretest and posttest. Twelve items, three from each of the four combinations of multiplication and division math operations, were randomly assigned to the pretest and the remaining twelve were assigned to the posttest. Thus, both the pretest and the posttest contained twelve two-step math word problems involving whole numbers. The problems on the pretest were in random order within each problem type and those on the posttest were in the same order by problem type. Each test consisted of three problems involving a multiplication operation followed by a second multiplication operation (multiply-multiply), three problems involving multiplication followed by division

(multiply-divide), three problems involving division followed by multiplication (divide-multiply), and three problems involving division followed by a second division operation (divide-divide).

The pretest was constructed and named as two different forms, Form A and Form B. The total of 24 items were randomly assigned as 12 pretest and 12 posttest items as described above. In Form A, problems one through six were written as non-personalized problems and problems seven through twelve were written as personalized problems. In Form B, the non-personalized problems fro m Form A (problems one to six on Form A) were converted to the personalized problems seven through twelve, and the personalized problems from Form A (problems seven through twelve on Form A) were converted to non-personalized problems one through six. Thus, the same 12 items appeared in both personalized and non-personalized forms across the two test versions, with the six items in non-personalized form as items 1-6 on each form and the six personalized items as items 7-12 on each form. The answer to each problem was scored as correct only when the correct final answer to the problem was given. The KR-20 reliability coefficient was .83 for the pretest.

Posttest. As with the pretest, posttest problems were developed and administered in two forms. The first six problems on Form A of the posttest were in non-personalized form and the second six problems were in personalized form. The first six problems on Form B of the posttest were items seven to twelve from Form A in non-personalized form and the final six problems on Form B (Items 7-12) were items one to six from Form A in personalized form. Thus, the problems on Form A and Form B of both the pretest and the posttest consisted of six non-personalized items followed by six personalized items. Like the answers on the pretest, each answer on the posttest was scored as correct or incorrect only based on the final answer. The KR-20 reliability coefficient was .87 for the posttest.

Subjects who received Form A on the pretest received Form B on the posttest, and those who received Form B on the pretest received Form A on the posttest. The overall mean proportions correct for all subjects across both tests were .68 for Form A and .69 for Form B.

Student Attitude Survey. A ten-item attitude survey served as the criterion measure for assessing the students' attitudes and motivation. Eight of the ten items were four-choice Likert-type questions that assessed student attitudes and continuing motivation toward the instruction. These items dealt with such matters as how interesting and how easy the instructional programs was, how much students learned from it, whether students could do two-step problems well, and whether they would like to do more problems like those in the program. Responses to these eight items were assigned a score of 4 for the most positive response and a score of 1 for the least positive response. The two remaining items were open-ended questions dealing with student likes and dislikes about the instructional program. The KR-20 reliability coefficient for the eight Likert-type items was .70.

Teacher Attitude Survey. A seven-item teacher survey, consisting of six four-choice Likert-type items and one open-ended question, was used to assess teacher attitudes toward the instructional program and the personalization strategy. Items dealt with topics such as the appropriateness of the program, its quality, whether it helped students learn, the effectiveness of personalization, and whether the teachers liked personalization. Teacher responses to the Likert-type items, like those of the students, were scored from 4 (most positive) to 1 (least positive).

Data Analysis

The data analysis for student achievement was a 2 (Treatment: Personalization and Non-personalization) x 2 (Ability Level: Higher Ability and Lower Ability) x 2 (Test Occasion: Pretest and Posttest) x 2 (Problem Type: Personalized and Non-personalized problems) repeated-measures univariate analysis of variance (ANOVA). Treatment and ability level were between-subjects variables in the analysis and test occasion and problem type were within-subjects variables. Attitude data were analyzed using a 2 (Treatment) x 2 (Ability Level) x 8 (Survey Items) multivariate analysis of variance (MANOVA) for the overall survey means.

Results

The results are reported in this section for achievement, student attitudes, and teacher attitudes.

Achievement

The pretest and posttest data for the two levels of personalization, the two ability levels, the two test occasions, and the two problem types are shown in Table 1. The achievement data for each variable in the four-factor design are discussed below.

Treatment. For personalization level, the mean overall scores across the two test occasions were 8.45 (70%) for the personalized subjects and 7.93 (66%) for the non-personalized subjects. ANOVA revealed that the difference in mean scores between the two personalization levels was not statistically significant, $F(1, 132) = 2.37$, $MSE = 3.80$, $p = .126$. However, the difference for treatment in this analysis was reduced due to the fact that the pretest and posttest scores were combined for each of the two levels of personalization because of the repeated-measures design. The mean scores by personalization level on the pretest (that is, prior to the treatment) were identical for the two levels: 6.54 items correct (55%) for the personalized treatment and 6.54 (55%) for the non-personalized treatment. In contrast, the mean posttest scores by treatment were 10.35 (86%) for the personalized treatment and 9.32 (78%) for the non-personalized treatment.

The differential pattern of identical scores for personalization level on the pretest, but a higher score for personalization than for non-personalization on the posttest, was reflected in a significant treatment by test occasion interaction, $F(1, 132) = 8.27$, $MSE = 1.09$, $p < .01$, $\eta^2 = .06$. A post hoc paired-samples t test revealed that the posttest score of 10.35 for the personalized treatment was significantly higher than the score of 9.32 for the non-personalized treatment, $t(67) = 3.01$, $p < .01$.

Table 1

Mean Scores by Treatment, Ability Level, Test Occasion, and Problem Type

	Problem Type			
	Personalized		Non-personalized	
Test Occasion and Treatment	Higher ability	Lower ability	Higher ability	Lower ability
Pretest				
Personalized				
M	4.79	1.94	4.71	1.65
SD	(0.98)	(1.28)	(1.09)	(1.28)
Non-personalized				
M	4.85	2.18	4.35	1.71
SD	(0.96)	(1.24)	(0.98)	(1.22)
Sub-totals				
M	4.82	2.06	4.53	1.68
SD	(0.96)	(1.26)	(1.04)	(1.24)
Posttest				
Personalized				
M	5.82	4.59	5.85	4.44
SD	(0.46)	(1.84)	(0.44)	(1.74)
Non-personalized				
M	5.68	3.94	5.35	3.68
SD	(0.59)	(1.69)	(0.95)	(1.80)
Sub-totals				
M	5.75	4.27	5.60	4.06
SD	(0.53)	(1.78)	(0.78)	(1.80)

Overall mean scores by test occasion and variables:

	Treatment	Ability	Problem Type	
Pretest Personalized	= 6.54	Higher = 9.35	Personalized	= 3.44
Non-personalized	= 6.54	Lower = 3.74	Non-personalized	= 3.10
Posttest Personalized	= 10.35	Higher = 11.35	Personalized	= 5.01
Non-personalized	= 9.32	Lower = 8.32	Non-personalized	= 4.83

Overall	Personalized	= 8.45	Higher = 10.35	Personalized	= 4.23
Non-personalized		= 7.93	Lower = 6.03	Non-personalized	= 3.97

Note. Total possible score equals 12 items correct on each test for treatment, ability level, and test occasion, and six items correct for each problem type.

Ability Level. Higher-ability students outscored lower-ability students across the two tests, 10.35 (86%) to 6.03 (50%). This difference for ability level was statistically significant, $F(1, 132) = 167.49$, $MSE = 3.80$, $p < .001$, $\eta^2 = .56$.

The 2 x 2 x 2 x 2 ANOVA also yielded a significant ability level by test occasion interaction, $F(1, 132) = 52.30$, $MSE = 1.09$, $p < .001$, $\eta^2 = .28$. This interaction reflected the fact that higher-ability students improved less from pretest to posttest than lower-ability students. Higher-ability subjects had mean scores of 9.35 (78%) on the pretest and 11.35 (95%) on the posttest, an improvement of 2.00 items correct, whereas lower-ability subjects had mean scores of 3.74 (31%) on the pretest and 8.32 (69%) on the posttest, an improvement of 4.58 items correct.

Test Occasion. The mean scores for test occasion were 6.54 (55%) for the pretest, and 9.84 (82%) for the posttest, a mean pretest-to-posttest increase of 3.30 items correct. This difference was statistically significant, $F(1, 132) = 338.85$, $MSE = 1.09$, $p < .001$, $\eta^2 = .72$. The two interactions that involved test occasion, treatment by test occasion and ability by test occasion, were reported above.

Problem Type. The overall mean scores for problem type were 4.23 (71%) for the six personalized problems and 3.97 (66%) for the six non-personalized problems, a statistically significant difference, $F(1, 132) = 17.63$, $MSE = .51$, $p < .001$, $\eta^2 = .12$. The ANOVA also yielded a significant two-way interaction for treatment by problem type, $F(1, 132) = 4.66$, $MSE = .51$, $p < .05$, $\eta^2 = .03$. This interaction reflected the fact that there was a rather small difference in the scores of subjects in the personalized treatment on the personalized and non-personalized problems (4.29, or 72%, on personalized items and 4.16, or 69%, on non-personalized items), and a larger difference in the scores of non-personalized subjects on these items (4.16, or 69%, on personalized items and 3.77, or 63%, on non-personalized items). Post hoc paired-samples t tests of the scores contributing to this interaction revealed that the mean score of 3.77 for non-personalized treatment on non-personalized items was significantly lower at the $p < .01$ level than the means of each of the other three groups.

Both personalized and non-personalized problems were included in the pretest in order to determine whether students would perform better on personalized word problems than on non-personalized problems prior to the instruction. The significant main-effect for problem type across test occasions indicates that subjects did perform better on personalized items than on non-personalized items both before and after they received instruction. Paired-sample t tests of the difference between scores on personalized and non-personalized items on the pretest and again on the posttest confirmed that significant achievement differences favoring personalized items occurred on both occasions.

Student Attitudes

The mean attitude scores by treatment and ability level for subjects' responses to the eight statements on the four-point Likert-type attitude survey administered after completion of the instructional program are shown in Table 2. Responses were scored as 4 for the most positive response to 1 for the most negative response.

The overall mean score across the eight Student Attitude Survey items was 3.42, a favorable rating indicating agreement with positive statements about the instructional program. The three highest-rated statements on the survey were "I learned a lot from this program" ($M = 3.71$), "It is important to know how to solve two-step math problems" ($M = 3.68$), and "I would like to do more math word problems like the ones in the program" ($M = 3.55$). The lowest-rated statement was "I am able to do two-step math problems well" ($M = 3.12$).

The data in Table 2 were analyzed using a 2 (treatment) x 2 (ability) x 8 (survey items) MANOVA to test for significant differences. A significant overall effect across the eight items was obtained for treatment, ($M = 3.52$ for personalization and 3.31 for non-personalization), $F(8, 127) = 7.10$, $MSE = .13$, $p < .001$, $\eta^2 = .08$, and for ability, ($M = 3.54$ for higher-ability students and 3.29 for lower-ability students), $F(8, 127) = 5.72$, $MSE = .13$, $p < .001$, $\eta^2 = .11$. The treatment by ability level interaction across the eight items was not statistically significant.

Univariate analyses on the eight survey items by personalization level revealed significantly more positive attitudes on four of the items for subjects in the personalized treatment than for those in the non-personalized treatment. Students in the personalized treatment had significantly more favorable scores at the $p < .001$ level on the items: "This program was interesting" ($M = 3.62$ for personalization and 3.24 for non-personalization, $\eta^2 = .09$),

"This program was easy" (M = 3.50 for personalization and 2.85 for non-personalization, η^2 = .19), and "This program had many familiar persons, places, and things" (M = 3.62 for personalization and 3.13 for non-personalization, η^2 = .11), and at the .05 level on the item "I would like to do more math word problems like the ones in the program" (M = 3.66 for personalization and 3.44 for non-personalization, η^2 = .04).

Univariate analyses for ability level revealed significant attitude differences favoring higher-ability students on five of the items. Higher-ability students responded significantly mo re positively than lower-ability subjects at the .001 level to the statements: "I learned a lot from this program" (M = 3.91 for higher ability and 3.51 for lower ability, η^2 = .14), and "I am able to do two-step math problems well" (M = 3.35 for higher ability and 2.88 for lower ability, η^2 = .11), and at the .05 level to the statements: "This program was easy" (M = 3.32 for higher ability and 3.03 for lower ability, η^2 = .04), "I liked this program" (M = 3.44 for higher ability and 3.12 for lower ability, η^2 = .05), and "I would like to do more math word problems like the ones in the program" (M = 3.68 for higher ability and 3.43 for lower ability, η^2 = .05).

The frequency of constructed responses on the attitude survey to the two open-ended questions about what students liked most and what they liked least was also tabulated. Student responses indicated that what they liked most was that the program was interesting, a response given by 42 of the 136 students (31%). The second most common response to what students liked most was the use of the names of cartoon characters, persons, and things with which they are familiar, a response indicated by 38 students (28%). When asked what they liked least about the program, 48 students (35%) indicated nothing and 33 students (24%) responded that the lesson was too difficult.

Table 2
Student Attitude Scores by Treatment and Ability Level

	Treatment			Ability		
Item	**Pers**	**Non-Pers**	**F**	**Higher**	**Lower**	**F**
1. It is important to know how to solve two-step math problems.	3.65	3.72	.57	3.72	3.65	.57
2. This program was interesting.	3.62	3.24	13.24**	3.40	3.46	.29
3. This program was easy.	3.50	2.85	31.00**	3.32	3.03	5.41*
4. This program had many familiar persons, places, and things.	3.62	3.13	15.81**	3.49	3.26	2.99
5. I learned a lot from this program.	3.69	3.74	.22	3.91	3.51	20.85**
I am able to do two-step math problems well.	3.16	3.07	.52	3.35	2.88	16.65**
I liked this program.	3.28	3.28	.00	3.44	3.12	6.29*
I would like to do more math word problems like the ones in the program.	3.66	3.44	4.81*	3.68	3.43	6.26*
Overall means	3.52	3.31	10.83**	3.54	3.29	16.00**

*$p < .05$. **$p < .01$.

Teacher Attitudes

The overall mean teacher rating on the six items on the Teacher Survey was 3.71, a favorable rating indicating strong agreement with positive statements about the instructional program. All four respondents agreed very strongly ($\underline{M}$ = 4.00) with three statements: "Personalization was a good teaching strategy", "I liked the personalized version of the program", and "I would enjoy teaching a personalized lesson to my students occasionally."

On the one open-ended question on the survey, "Please make any comments or suggestions that you would like to make about this instructional program," two teachers indicated that making math word problems more personalized increases students' motivation and interest. One teacher reported that the students would understand the personalized problems better because they could relate the information in the problems to their real-life situations and "shorten the distance" of their thinking patterns on the problems.

Discussion

The primary research question in this study addressed the issue of whether personalization of instruction increases the achievement of Taiwanese students on mathematics word problems. The treatment by test occasion interaction revealed that the personalized treatment did, in fact, result in significantly higher pretest-to-posttest gains than the non-personalized treatment. This positive finding for personalization of instruction is consistent with the results obtained in several studies of personalized mathematics instruction in the United States (Anand & Ross, 1987; Lopez & Sullivan, 1991, 1992). It differs from the overall result for personalization in the earlier study by Ku and Sullivan (2000) in Taiwan, in which a ceiling effect limited the potential gain of higher-ability fifth-graders.

The strategy of conducting this study with fourth-grade students, instead of fifth-graders as in the earlier study, had the desired result of reducing the strong ceiling effect observed previously (Ku & Sullivan, 2000). Despite the successful effort to reduce the ceiling effect with higher-ability students in this study, the significant ability level by test occasion interaction revealed that this effect was not completely eliminated. Higher-ability subjects averaged 95% (11.35 of 12 correct) on the posttest. This represented only 17% gain over their pretest score of 78%, but nevertheless approached the maximum possible score on the test. In contrast, lower-ability subjects were able to make a much higher gain, due in part to their much lower pretest score of only 31%.

The significant overall improvement from 55% on the pretest to 82% on the posttest indicated that the instructional program itself was generally effective across all treatment groups. This increase in scores occurred over an average time period of only about 53 minutes of instruction and review. In addition, the fact that students averaged 83% on the practice items and 82% on the posttest indicates that they retained their learning quite well from the instructional phase of the study to the final testing phase.

The finding that students scored higher on personalized test items than on non-personalized items, even prior to instruction (i.e., on the pretest), has potential implications for mathematics assessment. Certainly this finding merits further investigation in other settings and with larger samples to determine its consistency and generality. However, it suggests that students generally may score higher on math problems that include more familiar or personalized settings than on problems with unfamiliar or non-personalized settings. If that is the case, test developers who write standardized and criterion-referenced mathematics tests may want to consider the appropriateness of using settings for their problems that are generally familiar or popular with the projected test population.

The attitude data clearly indicated students' preference for personalized instruction, a result consistent with the findings in previous studies (Ku & Sullivan, 2000, Lopez & Sullivan, 1992, Ross & Anand, 1987). The strongest differences, all at the .001 level, were on items stating that the program was interesting; the program had many familiar persons, places, and things; and the program was easy. The first two of these statements reflect the intended nature of a personalized program and the third is consistent with the explanation that personalization may make learning easier by reducing subjects' cognitive load. Personalized subjects also agreed significantly more strongly ($p < .05$) that they would like to do more math word problems like the ones in the program, a statement that suggests greater continuing motivation or willingness to return-to-task on their part. In general, the significant differences on these particular items support the claimed motivational and learning advantages of personalization.

Several significant differences between ability levels on the attitude survey also reflected differences that one might expect between higher-ability and lower-ability students. Higher-ability students agreed more strongly that they learned a lot from the program ($p < .001$), that they were able to do two-step math problems well ($p < .001$), and that the program was easy ($p < .05$). These items appear to be the ones in the survey that were most

closely associated with ability level. In contrast, higher-ability and lower-ability subjects did not differ significantly in their responses to items that logically seem to be less associated with ability.

This study was not designed as a test of theoretical explanations for the effects of personalization. However, certain data from the teacher and student surveys tend to support the "increased interest" and "reduced cognitive load" explanations described in the introduction section of this paper.

The present results were obtained using group personalization in a low-technology environment that is typical of the schools in Taiwan and in many other countries. A basic use of personalization, or interest-based instruction, in such settings is for teachers to make a conscious effort to learn their students' interests and to incorporate them regularly into their instruction. Teachers can supplement their existing knowledge of student interests by using an interest survey, as was done in the present study, and/or by holding occasional discussions with their students about current topics and events that may be popular with them.

References

Anand, P. G., & Ross, S. M. (1987). Using computer-assisted instruction to personalize arithmetic materials for elementary school children. Journal of Educational Psychology, 79(1), 72-78.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. Journal of Educational Psychology, 88(4), 715-730.

D'Ailly, H. H., Simpson, J., & Mackinnon, G. E. (1997). Where should "you" go in a math compare problem? Journal of Educational Psychology, 89, 562-567.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. Journal of Educational Psychology, 83(1), 61-68.

Dwyer, H. J. (1996). Effect of personalization on reading comprehension. (Doctoral dissertation, Arizona State University, 1996). Dissertation Abstracts International, 59, Z3810.

Enright, B. E., & Choate, J. S. (1993). Mathematical problem solving: The goal of mathematics. In J. S. Choate (Eds.), Successful mainstreaming: Proven ways to detect and correct special needs (pp. 283). MA: Allyn and Bacon.

Hart, J. M. (1996). The effect of personalized word problems. Teaching Children Mathematics, 2(8), 504-505.

Herndon, J. N. (1987). Learner interests, achievement, and continuing motivation in instruction. Journal of Instructional Development, 10(3), 11-14.

Ku, H-Y, & Sullivan, H. J. (2000). Personalization of mathematics word problems in Taiwan. Educational Technology Research and Development, 48(3), 49-59.

Lopez, C. L., & Sullivan, H. J. (1991). Effects of personalized math instruction for Hispanic students. Contemporary Educational Psychology, 16(1), 95-100.

Lopez, C. L., & Sullivan, H. J. (1992). Effect of personalization of instructional context on the achievement and attitudes of Hispanic students. Educational Technology Research and Development, 40(4), 5-13.

Mayer, R. E. (1982). Memory for algebra story problems. Journal of Educational Psychology, 74(2), 199-216.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. Instructional Science, 26(2), 49-63.

Miller, D. C., & Kulhavy, R. W. (1991). Personalizing sentences and text. Contemporary Educational Psychology, 16(3), 287-292.

National Assessment of Educational Progress. (1992a). NAEP 1992 mathematics report card for the nation and the states. Washington, DC: National Center for Education Statistics. Report No. 23-ST02.

National Assessment of Educational Progress. (1992b). NAEP 1992 school effects on educational achievement in mathematics and science. Washington, DC: National Center for Education Statistics, Research and Development Report.

Rosen, D. R. (1984). Students' schemata for algebra word problems. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Ross, S. M. (1983). Increasing the meaningfulness of quantitative material by adapting context to student background. Journal of Educational Psychology, 75(4), 519-529.

Ross, S. M., & Anand, P. G. (1987). A computer-based strategy for personalizing verbal problems in teaching mathematics. Educational Communication and Technology Journal, 35(3), 151-162.

Ross, S. M., McCormick, D., & Krisak, N. (1986). Adapting the thematic context of mathematical problems to student interests: Individualized versus group-based strategies. Journal of Educational Research, 79(4), 245-252.

Ross, S. M., McCormick, D., Krisak, N., & Anand, P. G. (1985). Personalizing context in teaching mathematical concepts: Teacher-managed and computer-assisted models. Educational Communication and Technology Journal, 33(3), 169-178.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").